BRAIN
A JOURNAL OF NEUROLOGY

The dorsal stream contribution to phonological retrieval in object naming

Myrna F. Schwartz,[1] Olufunsho Faseyitan,[2] Junghoon Kim[1] and H. Branch Coslett[1,2]

1 Moss Rehabilitation Research Institute, Einstein Healthcare Network, Elkins Park, PA 19027, USA
2 Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA

Correspondence to: Myrna F. Schwartz, PhD,
Moss Rehabilitation Research Institute,
50 Township Line Road,
Elkins Park,
PA 19027,
USA
E-mail: mschwart@einstein.edu

Meaningful speech, as exemplified in object naming, calls on knowledge of the mappings between word meanings and phonological forms. Phonological errors in naming (e.g. GHOST named as 'goath') are commonly seen in persisting post-stroke aphasia and are thought to signal impairment in retrieval of phonological form information. We performed a voxel-based lesion-symptom mapping analysis of 1718 phonological naming errors collected from 106 individuals with diverse profiles of aphasia. Voxels in which lesion status correlated with phonological error rates localized to dorsal stream areas, in keeping with classical and contemporary brain-language models. Within the dorsal stream, the critical voxels were concentrated in premotor cortex, pre- and postcentral gyri and supramarginal gyrus with minimal extension into auditory-related posterior temporal and temporo-parietal cortices. This challenges the popular notion that error-free phonological retrieval requires guidance from sensory traces stored in posterior auditory regions and points instead to sensory-motor processes located further anterior in the dorsal stream. In a separate analysis, we compared the lesion maps for phonological and semantic errors and determined that there was no spatial overlap, demonstrating that the brain segregates phonological and semantic retrieval operations in word production.

Introduction

There is growing consensus that auditory language functions are divided across two cortical pathways (Scott and Johnsrude, 2003; Wise, 2003; Hickok and Poeppel, 2004, 2007; Saur et al., 2008; Ueno et al., 2011). From auditory processing regions in superior temporal gyri, the ventral route projects anteriorly and inferiorly in bilateral temporal lobes to effect the mapping of sound to meaning (Scott et al., 2000; Scott and Johnsrude, 2003; Crinion and

Price, 2005; Warren et al., 2009; Binder and Desai, 2011; Ueno et al., 2011). The dorsal pathway, more left-lateralized, is specialized for mapping sound to articulatory-based representations (Wernicke, 1874/1969; Warren et al., 2005; Pulvermüller et al., 2006; Saur et al., 2008). Hickok and Poeppel's (2004) influential dual stream model attributes the dorsal stream specialization to an auditory-motor integration circuit centred in the posterior Sylvian fissure at the parietal-temporal boundary, which they designate area Sylvian-Parietal-Temporal (Spt) (see Wise et al., 2001; and

Dhanjal *et al*., 2008, for evidence of polysensory functioning in this area). Over this circuit, information flows bi-directionally between auditory-phonological representations located in the posterior temporal cortex, and articulatory-motor speech representations located in prefrontal and central regions, with area Spt effecting the translation between sensory and motor codes.

It is proposed that the degree of involvement of the two streams in a particular language task depends on the type of mapping and the strategies by which the participant approaches the task. For example, under natural circumstances, perceiving and comprehending spoken language falls within the domain of the ventral route. However, laboratory speech perception tests, such as phonemic discrimination, frequently rely on covert speech and rehearsal, which engage dorsal stream circuitry (Hickok and Poeppel, 2004).

The dual stream model makes the non-obvious claim that a non-auditory, semantically driven speech task, such as object naming, relies on the dorsal route for error-free phonological production. As Buchsbaum *et al*. (2011) explain:

> There is reason to believe that speakers rely to some extent on an auditory-phonological memory of words they are attempting to produce, as Wernicke proposed, or in modern motor control terminology that the targets of a motor speech act are auditory in nature . . . If motor control of speech production is driven by auditory speech targets and if the link between auditory and motor systems is disrupted, one expects an increase in the error rate in speech production. (p. 125).

The dual stream model's hypothesized role for auditory-motor integration in semantically driven speech and naming piqued our interest for two reasons. First, it is in the spirit of the model to claim that the division of labour between the routes would shift to the ventral route in a task like object naming, which has a strong semantic component. Indeed, Hickok and Poeppel (2004) implied as much when they asserted that dorsal route involvement in naming and repetition is greatest under conditions of high phonologic load and low semantic constraint. Second, the auditory-motor integration mechanisms mentioned above (Buchsbaum *et al*., 2011) have largely been ignored in psycholinguistic models of naming, which focus instead on abstract representations and behavioural, rather than neural data. Thus, if their contribution can be definitively established, it could inspire models that better integrate cognitive and neural concepts. For important steps in this direction, see Hickok (2012) and Ueno *et al*. (2011). Seeking direct evidence that error-free phonological production in object naming depends on the dorsal route, and particularly on auditory-related posterior temporal and temporo-parietal cortices, the present study examined the anatomical basis of phonological errors in naming by mapping the lesions that correlate with phonological error rates in individuals with aphasia.

The dorsal route account of conduction aphasia

Arguably, the most compelling evidence that phonological processes in naming, along with those in repetition and short-term memory, rely on the dorsal route comes out of research on conduction aphasia. Patients who carry this diagnosis speak fluently and comprehend speech well, but they have limited phonological short-term memory, which impedes their ability to repeat. Additionally, in all types of production tasks (e.g. naming, as well as repetition), their speech is frequently marred by phonological errors (also called paraphasias) (Goodglass and Kaplan, 1983; Caplan *et al*., 1986).

The classical 19th century theory of aphasia attributes this syndrome to lesions of a white matter tract, the arcuate fasciculus, connecting Wernicke's area and Broca's area (Geschwind, 1965; Compston, 2006). Although the involvement of the arcuate fasciculus receives some support in contemporary investigations of conduction aphasia and its symptoms (Yamada *et al*., 2007; Duffau *et al*., 2008), other evidence suggests that the syndrome is actually caused by lesion in the nearby posterior superior temporal gyrus (Anderson *et al*., 1999; Quigg and Fountain, 1999; Hickok *et al*., 2000) or supramarginal gyrus (Axer *et al*., 2001; Baldo and Dronkers, 2006; Baldo *et al*., 2012). Hickok and Poeppel (2004) hypothesized that the repetition deficit and tendency to produce phonological errors might have a common basis in damage to area Spt and surrounding cortices. To test this hypothesis, Buchsbaum *et al*. (2011) performed an aggregate functional MRI analysis to map brain activity associated with auditory-motor integration processes in a verbal short-term working memory task, after which they performed a conjunction analysis of this functional MRI map with the normalized lesion distribution of 24 patients with conduction aphasia. The maximal lesion overlap circumscribed area Spt in the functional MRI map.

The current study

The dual stream model makes a clear prediction that phonological errors in naming should localize to the dorsal route auditory-motor circuit, particularly cortices in and surrounding area Spt. A goal of the present study was to test this prediction. We sought to improve on the available evidence in several ways. First, instead of mapping the lesions that give rise to the conduction aphasia syndrome, we mapped phonological errors themselves. Although discussions of phonological errors in aphasia tend to focus on conduction aphasia, such errors occur in most types of aphasia, especially during confrontation naming (e.g. Schwartz *et al*., 2006); it is generally accepted that common psychological and neurological mechanisms are involved (Dell *et al*., 1997; Fridriksson *et al*., 2009). By broadening our subject inclusion criteria, we were able to map the lesions that give rise to phonological errors in a large, diverse group of patients.

A second feature of our study is that all the phonological errors we mapped were produced in the context of picture naming. Critically, picture naming is a task in which access to phonology is driven by semantics, not auditory input. Therefore, this constitutes a strong test of the hypothesized involvement of posterior dorsal route circuitry in non-auditory speech tasks. The third feature is that we used voxel-based lesion-symptom mapping (VLSM) in a large cohort of patients with heterogeneous lesions and did not limit the analysis to predefined regions of interest.

The primary question we sought to answer was whether the VLSM of phonological errors would implicate cortices of the dorsal stream, the ventral stream or both. Additionally, within the dorsal stream, we were interested in discerning and possibly differentiating the contribution of posterior and anterior sectors. Table 1 lists the three alternative hypotheses.

The 'posterior dorsal stream' hypothesis embeds the Spt area prediction in a broader context, the focus of which is the role of the auditory-important posterior temporal and temporo-parietal cortices in accessing phonology for production. The supporting literature includes functional neuroimaging studies of word production and picture naming (Indefrey and Levelt, 2004; Graves et al., 2007; Wilson et al., 2009), as well as the studies of conduction aphasia discussed above. The 'anterior dorsal stream' hypothesis views phonological errors as fundamentally motor phenomena, caused by damage to parietal and frontal systems crucial for planning and programming motor speech (Guenther, 2006; Goldstein et al., 2007; Tian and Poeppel, 2010; Hickok, 2012; and Foundas et al., 1998; Cloutman et al., 2009, for evidence from neuropsychology). The 'ventral stream' hypothesis postulates that errors arise from faulty retrieval of lexical-phonological codes stored in or near semantically important areas of the temporal lobe. Candidate loci include superior and middle temporal cortices (Indefrey and Levelt, 2000, 2004; Graves et al., 2007; Wilson et al., 2010) and middle/inferior temporal and fusiform gyri (Raymer et al., 1997; Foundas et al., 1998; Hillis et al., 2002; DeLeon et al., 2007).

These anatomical hypotheses have potential relevance to an ongoing debate among language researchers regarding how much information sharing (i.e. interaction) takes place between semantic and phonological stages of lexical retrieval (Rapp and Goldrick, 2000). Guided by a theory that postulates limited interaction (Dell, 1986; Dell and O'Seaghdha, 1991), Dell and colleagues (1997) have argued that the computational deficits that give rise to phonological errors in aphasic naming are distinct from those that give rise to semantic and other lexical errors (Foygel and Dell, 2000; Schwartz et al., 2006). To begin to define the neural correlates of the respective computational operations, we carried out a VLSM analysis of semantic errors in naming based on 64 participants with chronic aphasia (Schwartz et al., 2009; Walker et al., 2011). In a region spanning the mid to anterior portion of the middle temporal gyrus and temporal pole, a cluster of voxels was identified that carried an association with semantic errors in naming, even after controlling for core semantic deficits and lesion size. In the present study, with a larger sample, we repeated the VLSM of semantic errors as performed by Schwartz et al. (2009) and compared it with the VLSM of phonological errors to determine whether there was spatial proximity or overlap among the voxels associated with these error types.

Materials and methods

Participants

To qualify for an ongoing project investigating the anatomical basis of psycholinguistic deficits in post-acute aphasia, participants were required to meet specified inclusion criteria, authorize release of relevant medical records and give informed consent to participate in a multi-session clinical and language assessment protocol approved by the Institutional Review Board at the Einstein Medical Centre. To determine the precise localization of their lesion, participants were asked to undergo structural MRI or CT brain imaging under a protocol approved by the Institutional Review Board at the University of Pennsylvania Medical School. In rare cases, a clinical imaging study was used to map the lesions, on condition that the scan was judged to be of high quality (i.e. high resolution and free from artefacts), and there had been no intervening change in neurological or cognitive status. Participants were paid for their participation and reimbursed for travel and related expenses.

The inclusion criteria were as follows: at least 1 month post-onset of aphasia secondary to stroke, living at home, medically stable without major psychiatric or neurological co-morbidities, premorbid right handedness, English as primary language, adequate vision and hearing (with or without correction) and CT or MRI confirmed left hemisphere cortical lesion. Of the 142 individuals enrolled, 106 qualified for the current study. The rest were excluded because they produced no correct responses in object naming ($n = 9$) or because their scans revealed bilateral damage or damage restricted to subcortical areas ($n = 27$).

The 106 study participants were 43% female and 46% African American. Their mean age was 58 years (range 26–79 years), and mean years of education was 14 years (range 10–21 years). Ninety-one (86%) were >6 months post-onset of aphasia; median months post-onset = 21.5; mean (SD) = 50.6 (66.6). Three were included who failed the hearing screen (Ventry and Weinstein, 1983) but had adequate conversational hearing. These three were not asked to perform the auditory input tasks in the language battery.

All participants performed a large assessment to determine their aphasia status, language skills and language-related cognitive functions. The assessment included the Western Aphasia Battery (Kertesz, 1982), Apraxia Battery for Adults (Dabul, 2000) and tests of word and non-word repetition, auditory lexical decision, verbal comprehension, short-term memory span and syntactic comprehension.

Table 1 Three alternative hypotheses

Hypothesis	Anatomical regions	Functional account
Posterior dorsal stream	pSTG, pSTS, Spt	Auditory-motor integration; auditory-phonological codes
Anterior dorsal stream	IFG, premotor, motor, precentral, SMG	Planning, programming motor speech
Ventral stream	Mid-posterior STG, mid-posterior MTG, temporal pole, pITG/fusiform	Semantic-phonological integration; lexical-phonological codes

pSTG = posterior superior temporal gyrus; pSTS = posterior superior temporal sulcus; Spt = Sylvian parieto-temporal area; IFG = inferior frontal gyrus; SMG = supramarginal gyrus; MTG = middle temporal gyrus; pITG = posterior inferior temporal gyrus.

Language tests and experimental measures

Philadelphia Naming test

The Philadelphia Naming Test (Roach *et al*., 1996) tests basic-level object naming with 175 pictured objects from a variety of semantic categories. Pictures have high familiarity, name agreement and image quality. Object names range in length from 1 to 4 syllables and in noun frequency from 1 to 2110 tokens per million (Francis and Kucera, 1982). The Philadelphia Naming Test was administered and scored according to standard procedures (http://mrri.org/philadelphia-naming-test). On each trial, the first complete (i.e. non-fragment) response produced within 20 s was scored and assigned to one of six response categories. A response was scored correct only if it exactly matched the designated target, unless the patient qualified as having apraxia of speech, meaning that his or her speech on the Apraxia Battery for Adults contained multiple instances of segmental distortion/substitution, together with abnormal prosody, slow rate (lengthened consonants and vowels) and/or sound, syllable or word prolongation often accompanied by intrusive schwa. In accordance with Philadelphia Naming Test scoring rules, patients with diagnosed apraxia of speech were allowed minor distortions that were consistent for that patient, and their naming responses that deviated from the target by the addition, deletion or substitution of a single consonant or consonant cluster were scored correct. All other responses were classified into error categories. The Philadelphia Naming Test error taxonomy and psycholinguistic rationale are discussed in earlier publications (Dell *et al*., 1997; Schwartz *et al*., 2006). The two error types that are most relevant to the present investigation are non-word errors (i.e. non-lexical paraphasias) and semantic errors.

Non-word errors are the prototypical errors of phonological access in production and the central focus of our study. Some taxonomic schemes reserve the phonological error code for responses that meet some threshold for phonemic overlap with the target. All such thresholds are arbitrary. The Philadelphia Naming Test scheme considers that a response is phonologically related to the target if the two share at least one phoneme in corresponding position or two phonemes in any position, excluding schwa. This phonological overlap criterion is used to classify certain errors and is optionally used to distinguish target-related non-words from target-unrelated non-words. Here, as in earlier work (Dell *et al*., 1997; Foygel and Dell, 2000; Schwartz *et al*., 2006), we opted to analyse all non-words, regardless of overlap. We based this decision on a theoretical model that views degree of target-error overlap as a continuous variable, reflecting the severity of a lexical-phonological access disorder and supporting evidence from longitudinal and cross-sectional studies in aphasia (Schwartz and Brecher, 2000; Schwartz *et al*., 2004). Thus, the corpus included errors such as (i) and (ii) below, which meet the Philadelphia Naming Test phonological overlap criterion, as well as those like (iii), which does not.

E.g.   (i) GHOST named as (→) 'goath'
     (ii) DINOSAUR → 'dinosaurus'
     (iii) APPLE → 'fuger'

Another model-informed decision was to exclude errors such as HORSE → house, where the response is phonologically (and not semantically) related to the target and comprises an actual word. Errors of this type, known as 'malapropisms' or 'formal errors', have a dual origin, according to the model. Although some result from faulty access to phonological units—the same mechanism that causes non-word errors—others originate at an earlier stage, where lexical-level units are selected (Gagnon *et al*., 1997; Schwartz *et al*., 2006).

In short, the analysis of phonological errors involved all and only Philadelphia Naming Test non-word errors. For each participant, these were expressed as a proportion of total trials ($n = 175$) to create the variable 'phonological errors'. For purposes of VLSM, phonological errors and the other analysed variables were transformed by square root to normalize the distribution.

The other Philadelphia Naming Test error type we analysed was semantic errors. These are real word responses that are semantically, and not phonologically, related to the target. A semantic relation, in this scheme, is a synonym, category coordinate, superordinate, subordinate or strong associate of the target (e.g. BOWL → 'vase'; FLOWER → 'rose'). For each participant, semantic errors were tabulated and expressed as a proportion of total errors to create the variable 'semantic errors'. We carried out a VLSM with semantic errors as the dependent variable to determine whether there was spatial proximity or overlap in the anatomical substrates for phonological and semantic errors.

Auditory Discrimination Test

In this 40-item test (Martin *et al*., 2005), the subject hears two recorded words ($n = 20$) or non-words ($n = 20$) in succession and indicates whether the two are the same or different. Non-identical pairs differ by a single onset or final phoneme. On half the trials, the second item immediately follows the first; on the remaining half, there is a filled 5-s delay (audible counting to five) between items, which encourages covert rehearsal during the delay. The number of errors proportional to trials constituted the analysed variable, 'auditory discrimination errors'. We carried out a VLSM with auditory discrimination errors as the dependent variable to see whether we could confirm the expected localization of auditory discrimination errors to the dorsal route auditory-motor circuit, and to ascertain the spatial proximity or overlap in the anatomical substrates for phonological errors and auditory discrimination errors. The three participants who failed the hearing screen were excluded from the VLSM of auditory discrimination errors.

Non-verbal comprehension tests

Two tests of non-verbal semantic comprehension were administered: the Pyramids and Palm Trees Test (Howard and Patterson, 1992) and the Camel and Cactus Test (Bozeat *et al*., 2000). Both involve picture–picture matching based on thematic relatedness (e.g. wine–grape). The 52-item Pyramids and Palm Trees Test requires a two-choice match; the 64-item Camel and Cactus Test requires a more demanding four-choice match. We averaged the square-root transformed error proportions to create a composite measure. In the VLSM of semantic error, this was entered as covariate of no interest to control for perceptual or core semantic deficits that might have resulted in a semantic naming error (Schwartz *et al*., 2009; Walker *et al*., 2011).

Lesion analysis

Image acquisition

Ninety-four patients received research MRI ($n = 56$) or CT ($n = 38$) brain scans at the Hospital of the University of Pennsylvania. High-resolution whole-brain T_1-weighted images (MPRAGE) were acquired for all the patients undergoing MRI. Of these, 49 were scanned on a 3-T Siemens Trio scanner (repetition time = 1620 ms, echo time = 3.87 ms, field of view = 192×256 mm, $1 \times 1 \times 1$-mm voxels). Seven patients had medical implants not approved for the higher strength magnetic field; they were scanned instead on a 1.5-T Siemens Sonata (repetition time = 3000 ms, echo time = 3.54 ms, field

of view = 24 cm, 1.25 × 1.25 × 1.2-mm voxels). For those patients who were not eligible for MRI scanning, whole-brain CT scans without contrast (60 axial slices, 3 mm thick) were acquired. Twelve additional patients declined scanning; for these patients, recent clinical scans (eight CT, four MRI) with clearly delineated lesion boundaries were substituted in the lesion tracing procedure.

Lesion segmentation and registration to template

Lesions imaged with MRI ($n = 60$) were segmented manually on a 1 × 1 × 1-mm T_1-weighted structural image by a trained technician who was blinded to the behavioural data. Lesion boundaries were identified on the basis of differences in signal characteristics between the infarcted and normal brain substance. Tissue with abnormal signal characteristics was included in the lesion; that is, 'lesions' included not only markedly abnormal (often cystic) regions but also regions surrounding the core area of infarction in which the signal characteristics differed from healthy grey or white matter. Efforts were made to define the contours of the lesions with precision. Areas reflecting secondary changes from the infarction (e.g. ventricular enlargement, regions with Wallerian degeneration) were not considered to be part of the lesion.

The structural scans were registered to a custom template, constructed from images acquired on the same scanner, using a symmetric diffeomorphic registration algorithm (Avants et al., 2006) implemented in the Advanced Normalization Tools package (http://picsl.upenn.edu/ANTS). Lesions were masked for warping using a variant of cost function masking (Brett et al., 2001). A single mapping from the custom intermediate template to the Montreal Neurological Institute (MNI) space 'Colin27' volume (Holmes et al., 1998) was used to complete the mapping from subject space to MNI space. This same mapping was then applied to the lesion maps. After being warped to MNI space, all of the manually drawn depictions of the lesions were compared with the original scan by investigator H.B.C., an experienced neurologist who was unfamiliar with the behavioural data.

Lesions imaged with CT ($n = 46$) were drawn by H.B.C. directly onto the Colin27 volume, after rotating (pitch only) the template to approximate the slice plane of the patient's scan. To this end, lesions were localized with respect to anatomical landmarks, such as sulci, gyri, deep grey matter structures, ventricles and the cortical surface. Using these landmarks, the lesion was rendered on the Colin27 volume. We have previously demonstrated excellent intra- and inter-rater reliability with this method (Schnur et al., 2009).

Lesion maps have intrinsic smoothness owing to the spatial coherence of the lesions (the estimated smoothness of our lesion maps is 6.4 × 7.0 × 5.1-mm full-width half-maximum). Consequently, we do not present analyses from smoothed data.

Voxel-based lesion-symptom mapping analyses

It is customary to exclude voxels from the analysis that have too few lesions to provide a stable comparison of lesioned/non-lesioned performance. The criterion is arbitrary. Here, we excluded voxels that were lesioned in <10 patients (~10% of the sample). For analyses with one dependent variable (e.g. phonological errors) and one independent variable (lesion status), a t-test was performed at each voxel comparing scores on the dependent variable between patients with and without lesions in that voxel. For analyses with one dependent variable and more than one independent variable (e.g. lesion status and a covariate of no interest), voxel-wise regression analyses were used instead. The VoxBo brain imaging package (http://www.nitrc.org/projects/voxbo/) was used to perform both types of analysis. The resulting t-maps were thresholded to control the false discovery rate (FDR) (Genovese et al., 2002) at $q = 0.01$, where q is the

expected proportion of false positives among supra-threshold voxels. The overlap analyses described in the 'Results' section used a more lenient FDR threshold ($q = 0.05$) to avoid underestimating the overlap of two statistical maps. For the two secondary VLSM analyses where we examined whether the lesion map for phonological errors would be affected by (i) regressing out lesion size or (ii) restricting the sample to participants without apraxia of speech, we adopted an arbitrary threshold of $t = 2.5$ because our goal was to determine whether these analyses would reveal any changes in the pattern of main results, rather than to establish brain areas significantly correlated with phonological errors.

The anatomical locations of voxels found to exceed threshold were determined by the judgment of H.B.C. in consultation with the Brodmann map and automated anatomical labelling atlas within MRIcro (www.mccauslandcenter.sc.edu/mricro/). The automated anatomical labelling atlas was used to locate and quantify the size of supra-threshold voxel clusters.

We were particularly interested in whether voxels associated with phonological errors would localize to area Spt. Area Spt has been defined in two ways. Hickok and Poeppel (2004; also Hickok, 2012) define Spt as the cortex in the posterior portion of the Sylvian fissure at the temporo-parietal junction. Area Spt may also, however, be delineated on the basis of functional MRI activations; defined in this manner, area Spt is most commonly located in the posterior portion of the planum temporale (Hickok et al, 2009; Buchsbaum et al., 2011). To explore the relationship between the brain regions identified in our VLSM study and area Spt, we generated a mask that includes both the posterior planum temporale and the junction of the parietal and temporal lobes (Fig. 2). As Hickok et al. (2009) and Buchsbaum et al. (2011) report substantial individual differences in the location of functionally defined area Spt, the mask included the full width of the planum temporale, extending from the lateral to the medial borders of the temporal lobe and extended from MNI $y = -23$ anteriorly to $y = -43$ posteriorly. The mask includes the site $(-51, -42, 21)$ reported by Buchsbaum et al. (2011) to represent area Spt.

Results

Behavioural findings

Aphasia severity, as measured by the Western Aphasia Battery aphasia quotient (Kertesz, 1982), ranged from 25.2 (moderately severe) to 97.9 (recovered); the mean was 73.3. The 10 participants who performed above the Western Aphasia Battery aphasia cut-off (>93.8) were diagnosed by an experienced speech-language therapist as having mild anomic aphasia based on extended speech samples elicited and analysed as part of the language battery. The presence of lengthy hesitations, repairs and omissions in these samples confirmed the participants' own experience of persisting word retrieval symptoms. For the rest of the participants, the standard Western Aphasia Battery criteria for subtype diagnosis were applied. The breakdown for the group overall was as follows: 49 participants (including the 10 'recovered') had anomic aphasia, 28 had Broca's aphasia, 17 had conduction aphasia, eight had Wernicke's aphasia, three had transcortical motor aphasia and one had global aphasia. Twenty-three participants (17 with Broca's aphasia) had apraxia of speech.

The total number of phonological errors across all participants was 1718; for those without apraxia of speech, it was 1095. Thus, nearly two-thirds of the phonological errors corpus came from patients without apraxia of speech. Eighty-one per cent of phonological errors qualified as target-related according to the Philadelphia Naming Test criterion (response and target shared at least one phoneme in corresponding position or two phonemes in any position, excluding schwa). Per item phonological errors rates correlated negatively with log lexical frequency (CELEX database; Baayen *et al.*, 1993) ($r = -0.52$; $P < 0.001$) and positively with phoneme length ($r = 0.79$; $P < 0.001$), confirming the known sensitivity of phonological errors to target length and frequency of usage (Dell, 1990; Nickels and Howard, 1995, 2004; Gordon, 2002; Kittredge *et al.*, 2008).

Table 2 shows proportion scores for phonological errors and other primary and derived Philadelphia Naming Test measures, including the breakdown for anomic, Broca, conduction and Wernicke groups. Mean phonological errors proportion was lowest in the anomic group and roughly comparable in the others. The ratio of phonological errors to the sum of phonological errors + semantic errors, indexing phonological errors specificity, was highest in the conduction group. However, both measures correlated with aphasia severity ($r = -0.47$ and 0.75 for proportion phonological errors and phonological errors specificity, respectively; both P's < 0.001); we ran separate ANOVAs with each measure to see whether there was a significant effect of group when severity was included as a covariate. There was not; for all effects involving group (main and interaction), F-values were <1.0. This justifies the inclusion of phonological error data from aphasics of all subtypes.

On the Auditory Discrimination Test, the mean auditory discrimination errors score was 0.15 (SD = 0.12), with range 0.00–0.50. Mean auditory discrimination errors by group were 0.10, 0.18, 0.19 and 0.27 for anomic, Broca, conduction and Wernicke groups, respectively. In an ANOVA that covaried for aphasia severity, main and interaction effects involving group did not approach significance (F-values < 1).

Lesion coverage

In VLSM, the power to detect brain–behaviour relationships at a given voxel depends on the number of patients with and without a lesion in the voxel. Maximal power is achieved in voxels lesioned in half the patients (53 in the present data set). The coverage map in Fig. 1 shows the overlap of lesions in all 106 patients. Coverage was excellent throughout the peri-Sylvian region, including posterior superior temporal gyrus (maximal voxel lesion count = 43), area Spt ($n = 42$), supramarginal gyrus ($n = 48$) and inferior frontal gyrus ($n = 51$). Within our ventral stream region of interest (Table 1), coverage ranged from good (maximal lesion counts 40 and 25 in middle temporal gyrus and temporal pole, respectively) to poor (counts <10 in some posterior inferior temporal gyrus/fusiform voxels). The inferolateral and mesial aspects of the temporal lobe lie outside the territory of the anterior circulation and therefore are commonly spared in patient samples such as this, which was selected for post-stroke aphasia. Thus, we were unable to explore the contributions of these areas of low coverage.

Anatomical findings: phonological errors

In the VLSM of phonological errors with FDR-controlled threshold ($q = 0.01$; critical t-value 3.51), a total of 14 372 significant voxels were identified (Fig. 2). The largest concentration was located in the postcentral gyrus (5331 voxels; 37% of total), in a region extending from the cortex abutting the Sylvian fissure inferiorly, to approximate MNI z-value 64, superiorly. The voxel with the highest t-value ($t = 5.17$) was located in the inferior portion of the precentral gyrus (MNI $x = -62$, $y = -16$, $z = 11$). There were also numerous significant voxels in the white matter deep to the precentral and, to a lesser degree, postcentral gyri. The AAL map designates the portions of the pre- and postcentral gyri overlying the insula as the 'Rolandic operculum'. There were 1419 significant voxels in this region. As their precise location is not available, these voxels are not included in the counts for the pre- and postcentral gyri. In this and subsequent analyses, the voxel counts for the pre- and postcentral gyri are underestimated. A large number of significant voxels (2275 voxels; 16% of total) occupied the supramarginal gyrus; the maximal t-value in this cluster ($t = 4.82$) was at MNI coordinates $x = -45$, $y = -30$, $z = 26$. A smaller cluster of significant voxels occupied the mid-portion of the central sulcus of the insula (1010 voxels; 7% of total); the maximal t-value ($t = 4.63$) was at MNI coordinates $x = -39$, $y = 2$, $z = -1$. Additionally, there was a small cluster in the inferior frontal gyrus pars opercularis and rare voxels in pars triangularis; the peak t-value ($t = 4.46$) in inferior frontal gyrus was at MNI coordinates $x = -61$, $y = -3$, $z = 4$.

Scattered suprathreshold voxels were found along the superior temporal gyrus on its most superior aspect. As these were contiguous with the large clusters in the postcentral gyrus and supramarginal gyrus, their presence in the superior temporal gyrus may have resulted from imprecision in the segmenting and/or warping of the lesions. Critical to the study hypotheses, there were no voxel clusters in the posterior superior temporal gyrus or posterior superior temporal sulcus. Nor was there much evidence of area Spt involvement. Only 39 voxels in the Spt mask correlated significantly with phonological errors—1% of the 3636 voxels in Spt mask and a negligible fraction of the 14 372 total voxels. Similar to what we saw in the superior temporal gyrus, the 39 Spt area voxels were concentrated on its superior border and contiguous with the large cluster in the supramarginal gyrus.

The overlap analyses used a more lenient FDR threshold ($q = 0.05$) to avoid underestimating overlap. Substituting this threshold in the VLSM of phonological errors increased the total significant voxels ($t \geqslant 2.41$) to 66 629 and produced an expansion of the central-region cluster anteriorly and posteriorly to encompass more of the precentral/premotor cortex and supramarginal gyrus, respectively (compare Fig. 2 with the area shown in blue in Figs 5, 6 and 7). At this reduced threshold, area Spt contained 684 significant voxels, 19% of voxels in the mask, 1% of total voxels.

Table 2 Philadelphia Naming Test scores

Participant group	Mean (SD)	Median	Lowest	Highest
Proportion correct				
All participants ($n = 106$)	0.64 (0.29)	0.74	0.01	0.98
Anomic ($n = 49$)	0.84 (0.09)	0.86	0.56	0.98
Broca's ($n = 28$)	0.47 (0.30)	0.55	0.01	0.89
Conduction ($n = 17$)	0.56 (0.25)	0.65	0.07	0.93
Wernicke's ($n = 8$)	0.26 (0.17)	0.21	0.07	0.58
Proportion phonological errors				
All participants	0.09 (0.10)	0.06	0.00	0.41
Anomic	0.04 (0.05)	0.03	0.00	0.30
Broca's	0.14 (0.10)	0.10	0.01	0.41
Conduction	0.14 (0.10)	0.13	0.01	0.31
Wernicke's	0.12 (0.11)	0.09	0.03	0.36
Proportion semantic errors				
All participants	0.03 (0.03)	0.03	0.00	0.12
Anomic	0.03 (0.02)	0.02	0.00	0.09
Broca's	0.04 (0.03)	0.03	0.01	0.11
Conduction	0.03 (0.02)	0.02	0.00	0.08
Wernicke's	0.05 (0.03)	0.05	0.01	0.12
Phonological errors/(semantic errors + phonological errors)				
All participants	0.62 (0.29)	0.73	0.00	1.00
Anomic	0.52 (0.32)	0.60	0.00	1.00
Broca's	0.68 (0.22)	0.71	0.31	0.99
Conduction	0.76 (0.22)	0.79	0.21	1.00
Wernicke's	0.64 (0.26)	0.72	0.20	0.94

The superior temporal gyrus contained 6015 voxels, most of which were again located along the superior aspect of the gyrus and were contiguous with the large supramarginal gyrus cluster. There were no voxel clusters in the posterior superior temporal gyrus or posterior superior temporal sulcus.

To summarize, in this, the main VLSM of phonological errors, we found that voxels carrying a statistically significant correlation with phonological errors were strongly concentrated in frontal and parietal cortices, i.e. the anterior dorsal stream. We carried out two secondary analyses to investigate possible confounding effects of lesion size and the inclusion of patients with apraxia of speech. In the first of these analyses, lesion size was entered as a covariate in a voxel-wise regression analysis predicting phonological errors. In the second, the original VLSM of phonological errors was re-run with the 23 patients with apraxia of speech excluded. We explored potential alterations in the basic pattern using an arbitrary threshold of $t = 2.5$. Voxels that were found to exceed this threshold were considered plausible (as distinct from statistically significant) carriers of a correlation with phonological errors.

Results are shown in Figs 3 and 4. Both secondary analyses again produced anterior-centred phonological errors-lesion correlation maps. In the analysis regressing out lesion size (Fig. 3), 13 596 voxels exceeded threshold. The largest concentrations were in the postcentral gyrus (4784 voxels; 35% of total) and supramarginal gyrus (2432 voxels; 18%), with a small cluster in the insula (949, 7%). In the VLSM of non-apraxic patients (Fig. 4), 14 377 voxels exceeded the threshold. Here, the largest concentration was in the supramarginal gyrus (4576; 32% of the total 14 377) rather than the postcentral gyrus (1665; 12%). Once again, there was extension into the superior temporal gyrus (3828 voxels; 27%) from the overlying frontal and parietal clusters; in this case, it included a small cluster anterior to Heschl's

gyri, contiguous with the voxel concentration in inferior frontal gyrus. The Spt mask contained 827 supra-threshold voxels (23% of the mask; 6% of total voxels).

Overlap with the arcuate fasciculus

In the VLSM of phonological errors, the presence of supra-threshold voxels in white matter underlying the cortex raises a question about the possible involvement of the arcuate fasciculus. Not having collected diffusion tensor imaging data from these participants, we were unable to address this directly. However, following Baldo *et al.* (2012), we performed an exploratory analysis based on a probabilistic diffusion tensor imaging map of the arcuate fasciculus obtained from the publicly available Johns Hopkins University Probabilistic Atlas (http://cmrm.med.jhmi.edu) (Zhang *et al.* 2010). A mask of arcuate fasciculus was constructed by thresholding the diffusion tensor imaging map at 0.8 probability. To calculate the extent of the overlap with the voxels associated with phonological errors, a second mask was built by thresholding the t-map with FDR controlled at $q = 0.05$. Of 10 582 voxels in the arcuate fasciculus mask, 29% (3077) overlapped with the t-map from the VLSM of phonological errors; the overlap was substantial in the anterior portion of the arcuate fasciculus (Fig. 5). The possibility that damage to the arcuate fasciculus is a sufficient or contributory factor in the genesis of phonological errors cannot be discounted and certainly warrants further investigation with more direct evidence (e.g. Duffau *et al.*, 2008; Saur *et al.*, 2008).

Auditory discrimination errors and its overlap with phonological errors

The VLSM of auditory discrimination errors with FDR controlled at $q = 0.05$ ($t = 2.72$) identified a total of 27 330 significant voxels. There was a large band extending from area Spt and the posterior superior temporal gyrus along the inferior supramarginal gyrus, inferior pre- and postcentral gyri, into the inferior frontal gyrus. Additional voxels in pre- and postcentral gyri extended superiorly to approximately MNI $z = -20$. In marked contrast to the pattern observed with phonological errors, a substantial number of voxels were found along the entire length of the superior temporal gyrus, with a distinct cluster in the posterior portion of the superior temporal sulcus. There were also numerous voxels in the posterior three-quarters of the planum temporale, including the voxel with maximal t-value (5.49). Area Spt was also heavily implicated; 2483 of the 3636 voxels in the mask (68%) were significant for auditory discrimination errors, accounting for 9% of total voxels in the auditory discrimination errors cluster. The extent to which auditory discrimination errors voxels overlapped the arcuate fasciculus (as defined above) was also assessed; the overlap was minimal.

Figure 6 presents the overlap between the auditory discrimination errors and phonological errors analyses; 15 015 voxels exceeded the threshold in both the phonological errors and auditory discrimination errors analyses. This set of shared voxels constitutes 55% of the total for auditory discrimination errors and 23% of the total for phonological errors. The great majority of shared voxels were in the pre- and postcentral gyri, as well as the
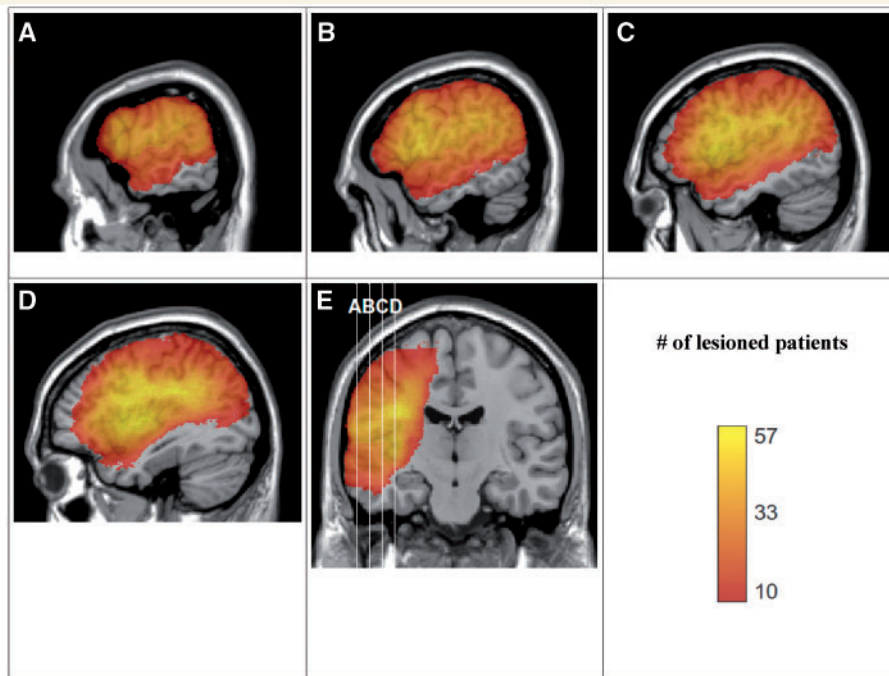
Figure 1 Maps depicting lesion overlaps of the 106 participants in the left hemisphere. Maps are superimposed on the MNI space Colin27 template. (**A–D**) MNI *x* coordinates: $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A single coronal slice at MNI *y* coordinate $= -16$ with white lines to indicate the location of sagittal slices **A–D** when viewed in the coronal plane. Voxels lesioned in at least 10 patients (the minimum allowed) are rendered in scale from red (10 patients) to yellow (57 patients).
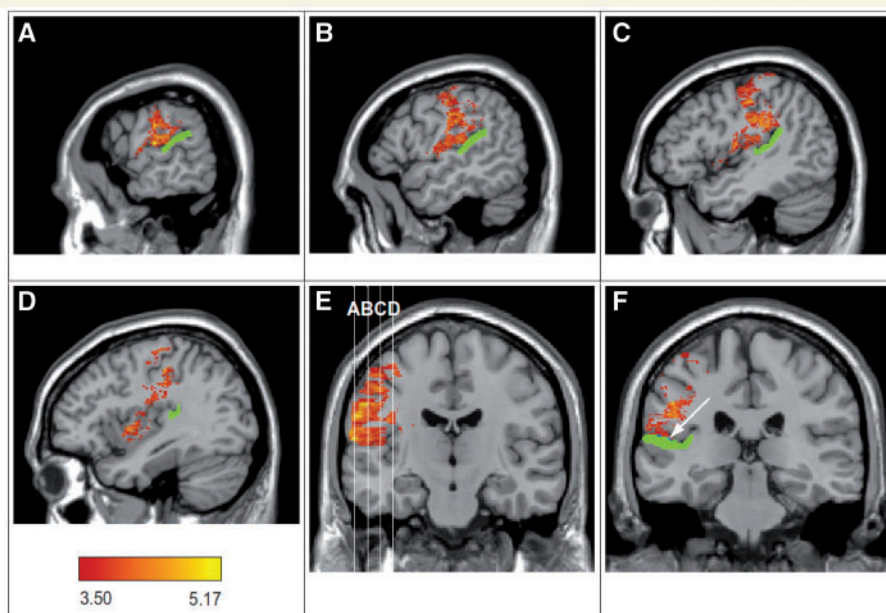
Figure 2 Statistical map (*t*-statistic) of phonological errors in picture naming between patients with and without lesion in each voxel. Map is thresholded with a FDR $q = 0.01$ ($t = 3.50$). Voxels rendered in red start at $t = 3.50$ and scale to the maximum *t*-value ($t = 5.17$) rendered in yellow. The light green area locates the Sylvian-parieto-temporal (Spt) area mask. Voxels in area Spt that exceeded the threshold are rendered in blue; in **F**, these are indicated by the arrow. (**A–D**) At MNI *x* coordinates: $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A coronal slice at MNI *y* coordinate -16, with white lines to indicate the location of sagittal slices **A–D** when viewed in the coronal plane. (**F**) A coronal slice at MNI *y* coordinate -28.
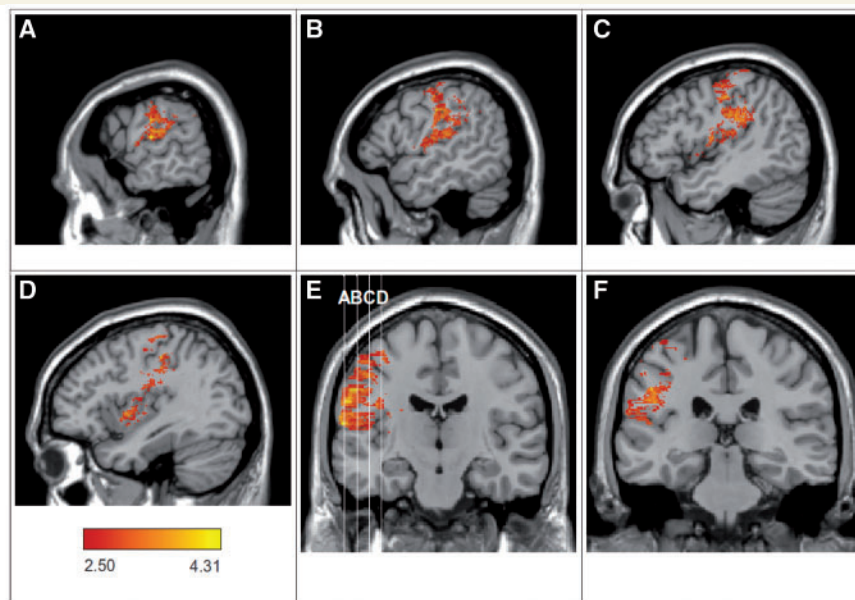
Figure 3 From the VLSM of phonological errors regressing out lesion volume. Map shows voxels thresholded at $t = 2.50$ and above displayed on the MNI-space Colin27 template. Voxels with the highest t-value (4.31) are rendered in yellow. (**A–D**) Sagittal slices at $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A single coronal slice at $y = -16$, with white lines indicating the location of slices **A–D** when viewed in the coronal plane. (**F**) A coronal slice at MNI y coordinate -28.
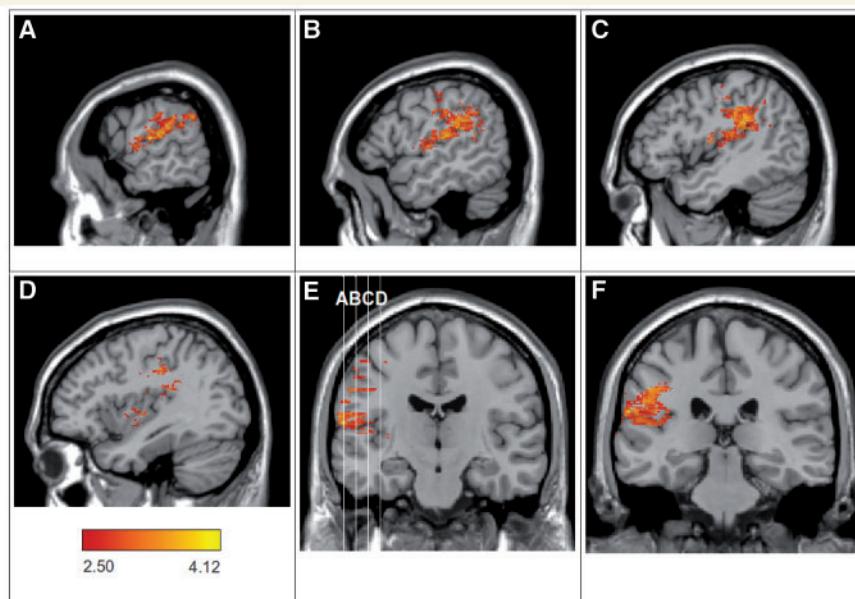
Figure 4 From the VLSM of phonological errors including only the 83 patients without apraxia of speech. Map shows voxels thresholded at $t = 2.50$ and above displayed on the MNI-space Colin27 template. Voxels with the highest t-value (4.12) are rendered in yellow. (**A–D**) Sagittal slices at $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A single coronal slice at $y = -16$, with white lines indicating the location of slices **A–D** when viewed in the coronal plane. (**F**) A coronal slice at MNI y coordinate -28.

supramarginal gyrus. There was also a small region of overlapping voxels in the posterior superior temporal gyrus, at the margin of the main cluster of voxels in the pre- and postcentral gyri. Additional small regions of overlapping voxels were identified in the middle superior temporal gyrus and inferior frontal gyrus [Brodmann area (BA) 44/45].

Semantic errors and its overlap with phonological errors

Finally, in Fig. 7, the voxels that exceeded threshold in the VLSM of semantic errors are presented together with those that were significant for phonological errors (FDR-corrected threshold,
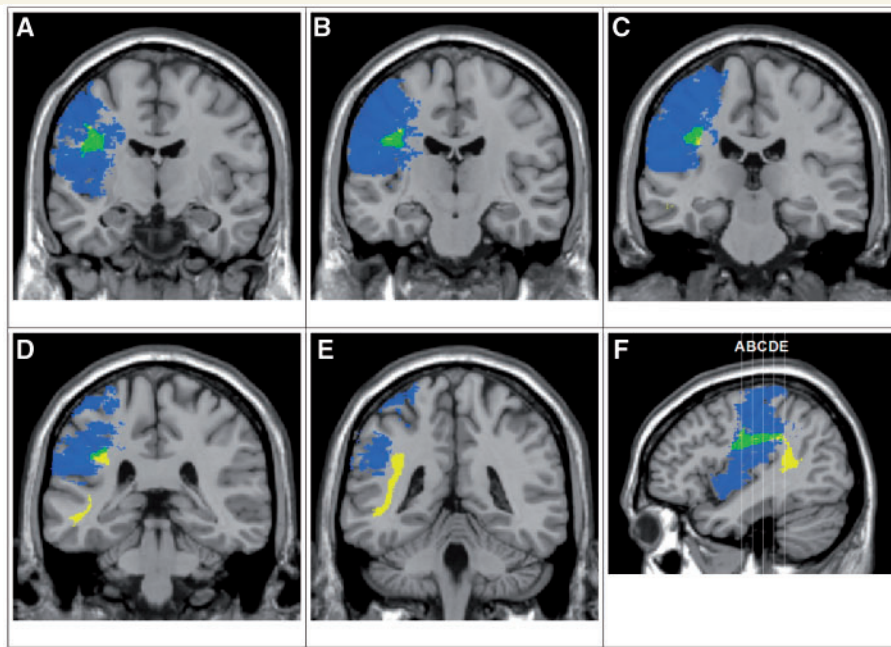
Figure 5 Map of the phonological errors results with a FDR $q = 0.05$ ($t = 2.41$) and a map of the arcuate fasciculus derived from the Johns Hopkins University Probabilistic atlas map. The probabilistic arcuate map is rendered at 0.8 probability. (**A–E**) MNI y coordinates: $y = -9$, $y = -17$, $y = -25$, $y = -33$ and $y = -41$, respectively. (**F**) A single sagittal slice at MNI x coordinate $= -40$ with white lines indicating the location of coronal slices (**A–D**) when viewed in the sagittal plane. Blue represents phonological errors results, yellow the white matter tracks of the arcuate fasciculus and green the region of overlap between phonological errors and the arcuate fasciculus.
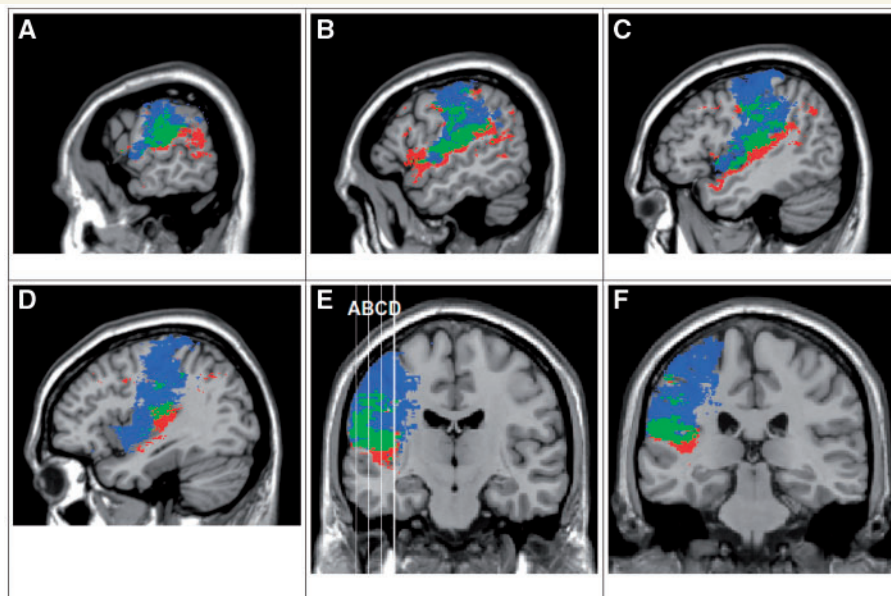

Figure 6 Lesions masks derived from the VLSM of auditory discrimination errors and the VLSM of phonological errors, thresholded with the same FDR correction ($q = 0.05$) are rendered together on the MNI-space Colin27 template. The critical t-value for auditory discrimination errors was 2.72; the critical t-value for phonological errors was 2.41. Overlapping voxels, i.e. those that surpassed the threshold in both analyses, are shown in green. Voxels that were significant in the auditory discrimination errors analysis only are shown in red; those that were significant in the phonological errors analysis only are shown in blue. (**A–D**) Sagittal slices at $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A single coronal slice at $y = -16$, with white lines indicating the location of slices **A–D** when viewed in the coronal plane. (**F**) A coronal slice at MNI y coordinate -28.
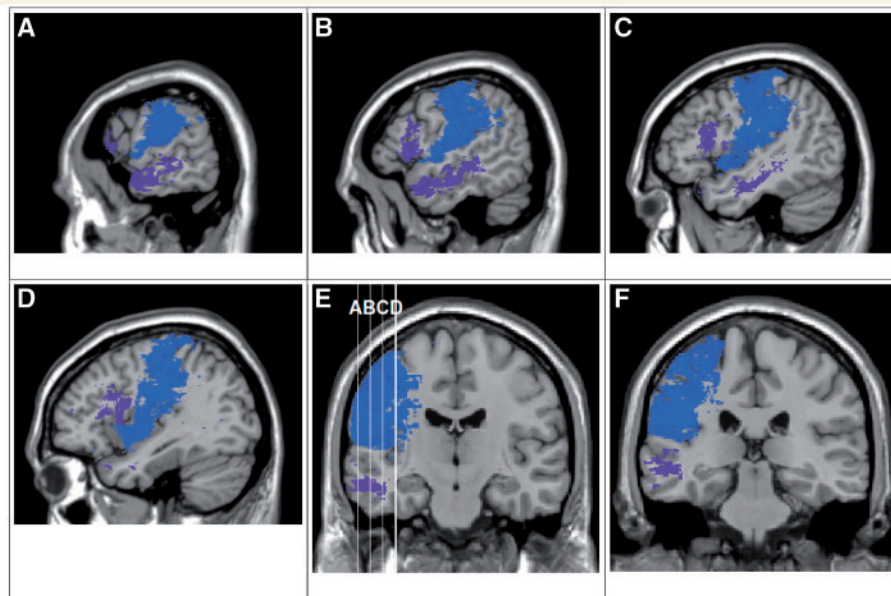
Figure 7 Lesions masks derived from the VLSM of semantic errors and the VLSM of phonological errors, thresholded with the same FDR correction ($q = 0.05$), are rendered together on the MNI-space Colin27 template. The critical t-value for semantic errors was 2.82; the critical t-value for phonological errors was 2.41. Voxels that were significant in the VLSM of semantic errors are shown in purple; those that were significant in the phonological errors analysis are shown in blue. No voxels were significant in both analyses. (**A–D**) Sagittal slices at $x = -62$, $x = -54$, $x = -46$ and $x = -38$, respectively. (**E**) A single coronal slice at $y = -16$, with white lines indicating the location of slices **A–D** when viewed in the coronal plane. (**F**) A coronal slice at MNI y coordinate -28.

$q = 0.05$ for both). As previously described (Schwartz et al., 2009; Walker et al., 2011), voxels associated with semantic errors were located in the mid-to-anterior portion of the middle temporal gyrus and the anterior portions of the inferior and middle frontal gyri. No voxels were associated with both semantic errors and phonological errors, i.e. the overlap was zero.

Discussion

This VLSM study provides definitive evidence that the dorsal system is essential for accurate phonological encoding of meaningful speech. The dual stream model claims that the dorsal route is specialized for mapping sound onto articulatory-based representations (i.e. auditory-motor integration), and that this specialization can be tapped by sub-lexical speech perception tasks (Hickok and Poeppel, 2004, 2007). Accordingly, we used the auditory discrimination task as a localizer for the dorsal route auditory-motor circuit and to determine its overlap with phonological errors. Fifty-five per cent of voxels that carried an association with auditory discrimination errors also carried an association with phonological errors. This would appear to confirm that disruption of dorsal stream processing is central to the explanation of phonological errors, even in confrontation naming, which is not an auditory input task. On the other hand, the explanatory emphasis on sound-based representations was not supported by the data. Of the two alternative dorsal stream hypotheses outlined in Table 1, the evidence strongly favoured the anterior version. The prominent involvement of frontal and parietal aspects of the

dorsal stream in the generation of phonological errors calls for a reinterpretation of the model's account of these errors.

Spatial segregation of phonological and semantic errors

The present data lend little support to the alternative we termed the 'ventral stream' hypothesis. Although there were voxels in mid to posterior region of superior temporal gyrus that correlated with phonological errors in the main and secondary analyses, the fact that most were scattered along the superior edge of the gyrus and continuous with the large, overlying frontal and parietal clusters, suggests to us that these were actually extensions of those clusters. Outside of the superior temporal gyrus, there were no voxels identified in the ventral stream. However, owing to the poor lesion coverage in the posterior inferior temporal gyrus/fusiform area, there was not sufficient power to detect effects there.

We used semantic errors in naming as a localizer for areas within the ventral system that participate in confrontation naming. Phonological errors voxels and semantic errors voxels did not overlap, despite the fact that both are manifestations of impairment by the same participants on the same task. Such spatial segregation of the neural substrates for phonological and semantic naming errors comports well with cognitive theories that assign semantic and phonological operations to distinct stages of lexical access (Garrett, 1975, 1980; Dell, 1986; Levelt et al., 1999).

Cloutman et al. (2009) also found evidence of spatial segregation in their study of the neural correlates of naming errors in

patients with acute left hemisphere stroke. Their participants underwent a language assessment within 24 h of the stroke, along with MRI to localize regions of tissue dysfunction (infarction or hypoperfusion). The study was primarily concerned with semantic naming errors, and a region of interest analysis showed that semantic error rate was predicted by tissue dysfunction in temporal lobe BA 22 and 37 (voxel-based analysis added evidence for prefrontal involvement.) Additionally, for patients who made >10% phonological errors in naming, a regional chi-square analysis showed that membership in this phonological error group was associated with tissue dysfunction in BA 6 (premotor cortex). Relevant to the semantic-phonological segregation question, phonological errors were not associated with any of the Brodmann areas that predicted semantic errors.

Fridriksson *et al*. (2009) conducted a functional imaging study that examined the functional anatomy of phonological and semantic errors in naming with 11 individuals with diverse aphasia profiles. They found that production of phonological errors produced activation in the left posterior perilesional occipital and temporal lobe areas. Production of semantic errors recruited approximately the same area, but in the right hemisphere (and see Postman-Caucheteux *et al*., 2010). In light of these findings, one might be tempted to conclude that the two types of errors have the same or overlapping structural basis (damage in the left occipital–temporal region) but differ as to whether the neural compensation takes place in perilesional tissue or in the right hemisphere homologue. On the contrary, our data provide strong evidence that the structural lesions that give rise to phonological and semantic errors do not localize to the same area.

Auditory-related posterior cortices

The VLSM of auditory discrimination errors revealed the expected involvement of dorsal stream circuitry, including robust effects in the posterior superior temporal gyrus and area Spt. This demonstrates that with a dependent variable that measures auditory perception and covert rehearsal, our patient sample and methods of analysis were capable of revealing effects in these posterior dorsal stream cortices. It is therefore not a function of low statistical power that in the VLSM of phonological errors, effects ranged from negligible to weak in area Spt, posterior superior temporal gyrus and posterior superior temporal sulcus, even when examined at the relaxed FDR-corrected threshold ($q = 0.05$) used to measure auditory discrimination errors/phonological errors overlap. The weak effects in these posterior regions, contrasted with the very strong effects in frontal and parietal regions, argue against the posterior dorsal stream hypothesis that auditory-important posterior cortices play an essential role in phonological retrieval in object naming.

A priori, there were good reasons to expect that the posterior dorsal stream hypothesis would be confirmed. For one thing, evidence from functional neuroimaging implicates the posterior superior temporal gyrus as a likely site of lexical-phonological form retrieval (Indefrey and Levelt, 2000, 2004). Graves *et al*. (2007) used overt picture naming and mapped voxels in which activation levels correlated with target word frequency, a lexical-phonological variable. In a second study, Graves *et al*. (2008) used repetition of non-words and mapped voxels in which activation was suppressed as a function of the number of times a non-word appeared as the target (i.e. phonological form learning). In both studies, posterior superior temporal gyrus activation correlated with the behavioural measures. Graves *et al*. (2008) reported a conjunction analysis with data from the two experiments, which identified an area in left posterior superior temporal gyrus (116 voxels; Talairach coordinates $-51 -38 +22$). The authors interpret this area as a substrate for phonological form retrieval in production. For related evidence, see Wilson *et al*. (2009).

Other functional neuroimaging studies have shown that the auditory cortices of the posterior superior temporal cortex play a joint role in perception and production. For example, Wise *et al*. (2001) identified a region in the left posterior temporal sulcus that activated both during heard speech and cued verbal recall. Buchsbaum *et al*. (2001) identified two foci within the posterior superior temporal gyrus that were active during both the perception and covert rehearsal phases of an auditory-verbal short-term memory task, one of which was at the termination of the Sylvian fissure (i.e. area Spt). Subsequent evidence demonstrated that area Spt also activated during perception and hummed rehearsal of tonal sequences (Hickok *et al*., 2003), and that lesions causing conduction aphasia overlapped with area Spt (Buchsbaum *et al*., 2011), all leading to the dual stream thesis that this temporoparietal region of the left hemisphere functions as an interface site for the integration of sensory and vocal tract-related motor representations in repetition, verbal short term memory and naming. The evidence regarding conduction aphasia provided an essential link in this chain of evidence, insofar as it suggested that lesions in area Spt were the cause of the phonological speech/naming errors that generally accompany the repetition and short-term memory deficits seen in these patients. However, here, in a direct investigation of the lesions that cause phonological errors in naming, we found no convincing evidence for the hypothesized role of area Spt in the production of such errors. Their association with other symptoms of conduction aphasia may be an accident of anatomical proximity to the areas further anterior in the dorsal stream that the present evidence identifies as the substrate for phonological errors.

Our evidence challenges the view that error-free phonological production in speech depends on auditory-motor integration processes in temporal-parietal cortices, but it does not rule out a role for auditory-motor integration mechanisms *per se*. For example, Ueno *et al*.'s (2011) neuro-computational model of the dual pathways assigns to inferior supramarginal gyrus—an area in which we found effects—the function of a hidden layer that extracts and represents the statistical structure shared between speech sounds and phonotactics. It is in the spirit of Ueno *et al*.'s (2011) model that damage to this hidden layer would result in phonological errors in speech and naming.

Regarding the evidence from Graves *et al*. (2007, 2008) and others demonstrating that the posterior superior temporal gyrus is a site of lexical-phonological retrieval in naming, in the face of that evidence, our negative results for this area may necessitate a modification of the view that phonological errors in speech derive from faulty access to lexical-phonological form information (e.g. Dell *et al*., 1997; Schwartz *et al*., 2004). One possibility is

that this psycholinguistic account underestimates the role of articulatory feedback in phonological retrieval (Plaut and Kello, 1999). A more radical possibility is that the units that enter into phonological errors are actually units of articulation. This suggestion comes out of modern accounts of phonology that view abstract phonology as an approximate or low-dimensional description of a complex dynamical sensory-motor system based on gestural units (Browman and Goldstein, 1992; Goldstein et al., 2007). Our findings can be seen as compatible with this articulatory phonology approach, and with some related approaches that build on motor-control models of speech. However, as we consider the anatomical evidence that points to a more grounded conceptualization of phonological errors, it is important to bear in mind that phonological error rates showed the expected correlation with target lexical frequency, and that the phonological errors map remained centred in the anterior dorsal stream even when patients with apraxia of speech were excluded from the analysis. Thus, it would be a mistake to think that the errors we analysed were not lexically inspired, or that the findings were unduly influenced by the inclusion of frank articulatory-motor errors.

The anterior dorsal stream hypothesis

We found that an association to phonological errors was carried by voxels in adjacent sectors of the supramarginal gyrus, postcentral, precentral and premotor cortices, i.e. closer to the motor side of the auditory-motor interface than auditory side. These areas are known to contribute to phonological processing in a wide variety of tasks (e.g. Vigneau et al., 2006; Rapcsak et al., 2009). The supramarginal gyrus features importantly in the literature on conduction aphasia and its attendant phonological deficits (Damasio and Damasio, 1980; Baldo and Dronkers, 2006; Fridriksson et al., 2010; Ueno et al., 2011; Baldo et al., 2012) and functional neuroimaging studies frequently reveal supramarginal gyrus activation in association with speech production and phonological processing (Shuster and Lemieux, 2005; Moser et al., 2009; Fridriksson et al., 2010; Hartwigsen et al., 2010). In a comprehensive meta-analysis of the literature on functional neuroimaging of language, Vigneau et al. (2006) identified five frontal and six temporal clusters engaged by phonological processing. Their frontal clusters correspond well to the fronto-central areas we identified (the supramarginal gyrus was counted among the temporal clusters.) Additional and more relevant corroboration of frontal involvement can be found in studies of the anatomical correlates of naming impairment in aphasia. Foundas et al. (1998) described a patient (Case 2) who presented acutely with fluent, anomic aphasia and a naming pattern that featured phonological errors along with circumlocutions. The patient had a lesion limited to inferior-lateral portions of BA 6 (premotor cortex). The group-level acute stroke study by Cloutman et al. (2009) also reported an association between BA 6 lesions and phonological errors in naming, as was described earlier.

Our evidence that the dorsal stream substrate for phonological errors centres on the fronto-central region and supramarginal gyrus resonates with motor-control models of speech production that highlight the role of somatosensory prediction and feedback (Tremblay et al., 2003). For example, Guenther's 'Directions into Velocities of Articulators' model of speech-motor control (Guenther, 2006; Guenther et al., 2006) incorporates auditory and somatosensory feedback control systems, with the somatosensory system centred on the postcentral gyrus and supramarginal gyrus. Building on the Directions into Velocities of Articulators model, Hickok's (2012) 'Hierarchical State Feedback Control' model associates somatosensory-based prediction and feedback with the phonemic level of psycholinguistic models, and auditory-based prediction and feedback with the syllabic level. As it is the phonemic level that is at play in phonological speech errors (Dell, 1986), these models of speech motor control suggest a possible explanation for why phonological errors are causally linked to lesions in postcentral gyrus and supramarginal gyrus.

Caveats and future directions

The anterior dorsal stream territory that carried effects for phonological errors spans multiple anatomical regions (premotor cortex, precentral gyrus, postcentral gyrus, supramarginal gyrus). We do not know which of these regions contribute jointly to the effect (lesions in both x and y are required), which contribute independently (lesions in either will produce the effect), and which arise spuriously from the spatial coherence of lesion data (neighbouring voxels/regions tend to be lesioned together). As discussed by Kimberg et al. (2007), the spatial coherence problem sets limits on what VLSM can tell about interregional differences: 'Even in the presence of good regional power, the power to detect differences between two regions may be low when the two regions are positively correlated' (Kimberg et al., 2007, p. 1075). As multivariate voxel-wise techniques become adapted for lesion data, a fuller picture of these interregional relationships may begin to emerge.

Integral to the fleshing out of interregional relationships is determining whether lesions affecting region x and those affecting y disrupt the same phonological process or different ones. For example, it may be that premotor lesions cause phonological errors by disrupting phonological processing at the motor planning stage, whereas supramarginal gyrus lesions do so by disrupting the selection or short-term buffering of phonological units. For behavioural patient studies addressing the causes of phonological errors, see Shallice et al. (2000), Schwartz et al. (2004), Romani and Galluzzi (2005) and Goldrick and Rapp (2007).

The present evidence may underestimate the contribution from areas outside the anterior dorsal stream territory. As previously noted, poor lesion coverage in the inferior temporal and fusiform gyri prevents us from addressing whether phonological errors in naming correlate with lesions in this basal temporal language area, which has been linked to phonological retrieval deficits in the naming performance of acute aphasics (e.g. De Leon et al., 2007). Another factor that reduces power to detect effects, even in regions with good lesion coverage, is anatomical variability. In particular, given the known interindividual variability in the location of area Spt, we cannot exclude the possibility that voxels in area Spt carried a correlation with phonological errors, but variability in their precise location resulted in insufficient overlap to identify with our univariate test. Analyses with multivariate techniques may shed light on this issue in the future.

The characterization of the lesions in this study likely underestimates the extent of functional tissue damage. In the future, this could be rectified by multimodal analyses that add information from perfusion-weighted and/or diffusion tensor MRI (e.g. Saur *et al.*, 2008; Cloutman *et al.*, 2009; Fridriksson *et al.*, 2010). Functional neuroimaging will also be useful in determining which posterior dorsal stream and ventral stream regions show reduced function as a consequence of lesions localized to the anterior dorsal stream (Crinion *et al.*, 2006).

Conclusion

The present study, the largest lesion analysis of phonological errors ever conducted, provides compelling new evidence on the anatomical basis of phonological errors in naming. The data confirm that although the naming task is driven by semantics, these naming errors arise from disruption of dorsal and not ventral stream processes. Indeed, perhaps the most striking of our findings was the segregation of the neural substrates for phonological and semantic errors in naming (Fig. 7). This bolsters the 'primary systems' view that phonological and semantic systems (along with syntax) are primary and distinct (Dell, 1986; Plaut *et al.*, 1996). From a clinical perspective, it adds confidence that aphasia diagnostic schemes built on these pillars offer a sound alternative to the classical Broca-Wernicke scheme (Mesulam *et al.*, 2009; Henry *et al.*, 2012).

Of the three hypothesized accounts of the genesis of phonological errors, the evidence strongly favours the anterior dorsal stream hypothesis. The association with phonological errors localized to premotor cortex, pre-and postcentral gyri and supramarginal gyrus, leading us to speculate that for naming, at least, the cause of phonological errors may relate less to auditory guidance of speech than to its motor planning and programming.

Acknowledgements

Funding

References

Anderson JM, Gilmore R, Roper S, Crosson B, Bauer RM, Nadeau S, et al. Conduction aphasia and the arcuate fasciculus: a reexamination of the Wernicke-Geschwind model. Brain Lang 1999; 70: 1–12.

Avants B, Schoenemann PT, Gee JC. Lagrangian frame diffeomorphic image registration: morphometric comparison of human and chimpanzee cortex. Med Image Anal 2006; 10: 397–412.

Axer H, von Keyserlingk AG, Berks G, von Keyserlingk DG. Supra- and infrasylvian conduction aphasia. Brain Lang 2001; 76: 317–31.

Baayen RH, Piepenbrock R, Van Rijin H. The CELEX lexical database. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium; 1993.

Baldo JV, Dronkers NF. The role of inferior parietal and inferior frontal cortex in working memory. Neuropsychology 2006; 20: 529–38.

Baldo JV, Katseff S, Dronkers NF. Brain regions underlying repetition and auditory-verbal short-term memory deficits in aphasia: evidence from voxel-based lesion symptom mapping. Aphasiology 2012; 26: 338–54.

Binder JR, Desai RH. The neurobiology of semantic memory. Trends Cogn Sci 2011; 15: 527–36.

Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, Hodges JR. Non-verbal semantic impairment in semantic dementia. Neuropsychologia 2000; 38: 1207–15.

Brett M, Leff AP, Rorden C, Ashburner J. Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage 2001; 14: 486–500.

Browman CP, Goldstein L. Articulatory phonology: an overview Haskins laboratories status report on speech research 1992; 111 (112): 23–42.

Buchsbaum BR, Baldo JV, Okada K, Berman KF, Dronkers N, D'Esposito M, et al. Conduction aphasia, sensory-motor integration, and phonological short-term memory—an aggregate analysis of lesion and fMRI data. Brain Lang 2011; 119: 119–28.

Buchsbaum BR, Hickok G, Humphries C. Role of the left posterior superior temporal gyrus in phonological processing for speech perception and production. Cogn Sci 2001; 25: 663–78.

Caplan D, Vanier M, Baker E. A case study of reproduction conduction aphasia: 1. Word production. Cogn Neuropsychol 1986; 3: 99–128.

Cloutman L, Gottesman R, Chaudhry P, Davis C, Kleinman JT, Pawlak M, et al. Where (in the brain) do semantic errors come from? Cortex 2009; 45: 641–9.

Compston A. From the archives. Brain 2006; 129: 1347–50.

Crinion J, Price CJ. Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. Brain 2005; 128: 2858–71.

Crinion JT, Warburton EA, Lambon Ralph MA, Howard D, Wise RJS. Listening to narrative speech after aphasic stroke: the role of the left anterior temporal lobe. Cereb Cortex 2006; 16: 1116–25.

Dabul BL. Apraxia battery for adults, second edition. Austin, TX: Pro-Ed; 2000.

Damasio A, Damasio H. The anatomical basis of conduction aphasia. Brain 1980; 103: 337–50.

DeLeon J, Gottesman RF, Kleinman JT, Newhart M, Davis C, Heidler-Gary J, et al. Neural regions essential for distinct cognitive processes underlying picture naming. Brain 2007; 130: 1408–22.

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychol Rev 1986; 93: 283–321.

Dell GS. Effects of frequency and vocabulary type on phonological speech errors. Lang Cogn Proc 1990; 5: 313–49.

Dell GS, O'Seaghdha PG. Mediated and convergent lexical priming in language production: a comment on Levelt et al (1991). Psychol Rev 1991; 98: 604–14.

Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. Psychol Rev 1997; 104: 801–38.

Dhanjal NS, Handunnetthi L, Patel MC, Wise RJ. Perceptual systems controlling speech production. J Neurosci 2008; 28: 9969–75.

Duffau H, Gatignol P, Mandonnet E, Capelle L, Taillandier L. Intraoperative subcortical stimulation mapping of language pathways in a consecutive series of 115 patients with Grade II glioma in the left dominant hemisphere. J Neurosurg 2008; 109: 461–71.

Foundas A, Daniels SK, Vasterling JJ. Anomia: case studies with lesion localization. Neurocase 1998; 4: 35–43.

Foygel D, Dell GS. Models of impaired lexical access in speech production. J Mem Lang 2000; 43: 182–216.

Francis WN, Kucera H. Frequency analysis of English usage: lexicon and grammar. Boston, MA: Houghton Mifflin; 1982.

Fridriksson J, Baker JM, Moser D. Cortical mapping of naming errors in aphasia. Hum Brain Mapp 2009; 30: 2487–98.

Fridriksson J, Kiartansson O, Morgan PS, Hialtason H, Magnusdottir S, Bonilha L, et al. Imparied speech repetition and left parietal lobe damage. J Neurosci 2010; 30: 11057–61.

Gagnon DA, Schwartz MF, Martin N, Dell GS, Saffran EM. The origins of formal paraphasias in aphasics' picture naming. Brain Lang 1997; 59: 450–72.

Garrett MF. The analysis of sentence production. In: Bower GH, editor. The psychology of learning and motivation. London: Academic Press; 1975. p. 133–75.

Garrett MF. Levels of processing in sentence production. In: Butterworth B, editor. Language production. London: Academic Press; 1980. p. 177–220.

Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 2002; 15: 870–8.

Geschwind N. Disconnection syndromes in animals and man. Part II. Brain 1965; 88: 585–644.

Goldrick M, Rapp B. Lexical and post-lexical phonological representations in spoken production. Cognition 2007; 102: 219–60.

Goldstein L, Pouplier M, Chen L, Saltzman E, Byrd D. Dynamic action units slip in speech production errors. Cognition 2007; 103: 386–412.

Goodglass H, Kaplan E. The assessment of aphasia and related disorders. 2nd edn. Philadelphia, PA: Lea & Febiger; 1983.

Gordon JK. Phonological neighborhood effects in aphasic speech errors: spontaneous and structured contexts. Brain Lang 2002; 82: 113–45.

Graves WW, Grabowski TJ, Mehta S, Gordon JK. A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. J Cogn Neurosci 2007; 19: 617–31.

Graves WW, Grabowski TJ, Mehta S, Gupta P. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. J Cogn Neurosci 2008; 20: 1698–710.

Guenther FH. Cortical interactions underlying the production of speech sounds. J Commun Disord 2006; 39: 350–65.

Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 2006; 96: 280–301.

Hartwigsen G, Baumgaertner A, Price CJ, Koehnke M, Ulmer S, Siebner HR. Phonological decisions require both the left and right supramarginal gyri. Proc Natl Acad Sci USA 2010; 107: 16494–9.

Henry ML, Beeson PM, Alexander GE, Rapcsak SZ. Written language impairments in primary progressive aphasia: a reflection of damage to central semantic and phonological processes. J Cogn Neurosci 2012; 24: 261–75.

Hickok G. Computational neuroanatomy of speech production. Nat Rev Neurosci 2012; 13: 135–45.

Hickok G, Buchsbaum BR, Humphries C, Muftuler T. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Sylvian parieto-temporal area. J Cogn Neurosci 2003; 15: 673–82.

Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, et al. A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. Neurosci Lett 2000; 287: 156–60.

Hickok G, Okada K, Serences JT. Area Sylvian parieto-temporal area in the human planum temporale supports sensory-motor integration for speech processing. J Neurophysiol 2009; 101: 2725–32.

Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 2004; 92: 67–99.

Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci 2007; 8: 393–402.

Hillis AE, Kane A, Tuffiash E, Ulatowski JA, Barker P, Beauchamp NJ, et al. Reperfusion of specific brain regions by raising blood pressure restores selective language functions in subacute stroke. Brain Lang 2002; 79: 495–510.

Holmes CJ, Hoge R, Collins L, Woods R, Toga AW, Evans AC. Enhancement of MR images using registration for signal averaging. J Comput Assist Tomogr 1998; 22: 324–33.

Howard D, Patterson K. Pyramids and palm trees: a test of semantic access from pictures and words. Bury St. Edmunds, UK: Thames Valley Test Company; 1992.

Indefrey P, Levelt WJM. The neural correlates of language production. In: Gazzaniga MS, editor. The new cognitive neurosciences. 2nd edn. Cambridge, MA: MIT Press; 2000.

Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. Cognition 2004; 92: 101–44.

Kertesz A. Western aphasia battery test manual. 2nd edn. New York, NY: Grune & Stratton; 1982.

Kimberg DY, Coslett HB, Schwartz MF. Power in boxel-based lesion-symptom mapping. J Cogn Neurosci 2007; 19: 1067–80.

Kittredge AK, Dell GS, Verkuilen J, Schwartz MF. Where is the effect of frequency in word production? Insights from aphasic picture naming errors. Cogn Neuropsychol 2008; 25: 463–92.

Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behav Brain Sci 1999; 22: 1–75.

Martin N, Schwartz MF, Kohen FP. Assessment of the ability to process semantic and phonological aspects of words in aphasia: a multi-measurement approach. Aphasiology 2005; 20: 1–13.

Mesulam M, Wieneke C, Rogalski E, Cobia D, Thompson C, Weintraub S. Quantitative template for subtyping primary progressive aphasia. Arch Neurol 2009; 66: 1545–51.

Moser D, Baker JM, Sanchez CE, Rorden C, Fridriksson J. Temporal order processing of syllables in the left parietal lobe. J Neurosci 2009; 29: 12568–73.

Nickels L, Howard D. Aphasic naming: what matters? Neuropsychologia 1995; 33: 1281–303.

Nickels L, Howard D. Dissociating effects of number of phonemes, number of syllables, and syllabic complexity on word production in aphasia: it's the number of phonemes that counts. Cogn Neuropsychol 2004; 21: 57–78.

Plaut DC, Kello CT. The emergence of phonology from the interplay of speech comprehension and production: a distributed connectionist approach. In: MacWhinney B, editor. The emergence of language. Mahwah, NJ: Lawrence Erlbaum Associates Inc.; 1999. p. 381–415.

Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: computational principles in quasi-regular domains. Psychol Rev 1996; 103: 56–115.

Postman-Caucheteux WA, Birn RM, Pursley RH, Butman JA, Solomon JM, Picchioni D, et al. Single-trial fMRI shows contralesional activity linked to overt naming erors in chronic aphasic patients. J Cogn Neurosci 2010; 22: 1299–318.

Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F. Motor cortex maps articulatory features of speech sounds. Proc Natl Acad Sci 2006; 103: 7865–70.

Quigg M, Fountain NB. Conduction aphasia elicited by stimulation of the left posterior superior temporal gyrus. J Neurol Neurosurg Psychiatry 1999; 66: 393–6.

Rapcsak SZ, Beeson P, Henry ML, Leyden A, Kim E, Rising K, et al. Phonological dyslexia and dysgraphia: cognitive mechanisms and neural substrates. Cortex 2009; 45: 575–91.

Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. Psychol Rev 2000; 107: 460–99.

Raymer AM, Maher LM, Foundas AL, Heilman KM, Rothi LJG. Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia. Brain Cogn 1997; 34: 287–92.

Roach A, Schwartz MF, Martin N, Grewal RS, Brecher A. The Philadelphia naming test: scoring and rationale. Clin Aphasiol 1996; 24: 121–33.

Romani C, Galluzzi C. Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. Cogn Neuropsychol 2005; 22: 817–50.

Saur D, Kreherb W, Schnellb S, Kümmerer D, Kellmeyer P, Vrya M-S, et al. Ventral and dorsal pathways for language. Proc Natl Acad Sci 2008; 105: 18035–40.

Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. Proc Natl Acad Sci 2009; 106: 322–7.

Schwartz MF, Brecher A. A model-driven analysis of severity, response characteristics, and partial recovery in aphasics' picture naming. Brain Lang 2000; 73: 62–91.

Schwartz MF, Dell GS, Martin N, Gahl S, Sobel P. A case-series test of the interactive two-step model of lexical access: evidence from picture naming. J Mem Lang 2006; 54: 228–64.

Schwartz MF, Kimberg DY, Walker GM, Faseyitan O, Brecher A, Dell GS, et al. Anterior temporal involvement in semantic word retrieval: VLSM evidence from aphasia. Brain 2009; 132: 3411–27.

Schwartz MF, Wilshire CE, Gagnon DA, Polansky M. Origins of nonword phonological errors in aphasic picture naming. Cogn Neuropsychol 2004; 21: 159–86.

Scott SK, Blank CC, Rosen S, Wise RJS. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 2000; 123: 2400–6.

Scott SK, Johnsrude IS. The neuroanatomical and functional organization of speech perception. Trends Neurosci 2003; 26: 100–7.

Shallice T, Rumiati RI, Zadini A. The selective impairment of the phonological output buffer. Cogn Neuropsychol 2000; 17: 517–46.

Shuster LI, Lemieux SK. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang 2005; 93: 20–31.

Tian X, Poeppel D. Mental imagery of speech and movement implicates the dynamics of internal forward models. Front Psychol 2010; 1: 1–23.

Tremblay S, Shiller DM, Ostry DJ. Somatosensory basis of speech production. Nature 2003; 423: 866–9.

Ueno T, Saito S, Rogers TT, Lambon Ralph MA. Lichtheim 2: synthesising aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. Neuron 2011; 72: 385–96.

Ventry IM, Weinstein BE. Identification of elderly people with hearing problems. ASHA 1983; 25: 37–42.

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, et al. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. Neuroimage 2006; 30: 1414–32.

Walker GM, Schwartz MF, Kimberg DY, Faseyitan O, Brecher A, Dell GS, et al. Support for anterior temporal involvement in semantic error production in aphasia: new evidence from VLSM. Brain Lang 2011; 117: 110–22.

Warren JE, Crinion JT, Lambon Ralph MA, Wise RJS. Anterior temporal lobe connectivity correlates with functional outcome after aphasic stroke. Brain 2009; 132: 3428–42.

Warren JE, Wise RJS, Warren JD. Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci 2005; 28: 636–43.

Wernicke C. The aphasic symptom complex: a psychological study on a neurological basis. Breslau: Cohn and Weigert. Reprinted. In: Cohen RS, Wartofsky MW, editors. Boston studies in the philosophy of science. Vol. 4. Boston, MA: Reidel; 1874/1969.

Wilson SM, Henry ML, Besbris M, Ogar JM, Dronkers NF, Jarrold W, et al. Connected speech production in three variants of primary progressive aphasia. Brain 2010; 133: 2069–88.

Wilson SM, Isenberg AL, Hickok G. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. Hum Brain Mapp 2009; 30: 3596–608.

Wise RJS. Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. Br Med Bull 2003; 65: 95–119.

Wise RJS, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural subsystems within 'Wernicke's area'. Brain 2001; 124: 83–95.

Yamada K, Nagakane Y, Mizumo T, Hosomi A, Nakagawa M, Nishimura T. MR tractography depicting damage to the arcuate fasciculus in a patient with conduction aphasia. Neurology 2007; 68: 789.

Zhang Y, Zhang J, Oishi K, Faria AV, Jiang H, Li X, et al. Atlas-guided tract reconstruction for automated and comprehensive examination of the white matter anatomy. Neuroimage 2010; 52: 1289–301.